(12) DEMANDE INTERNATIONALE PUBLIÉE EN VERTU DU TRAITÉ DE COOPÉRATION EN MATIÈRE DE BREVETS (PCT)

(19) Organisation Mondiale de la Propriété Intellectuelle
Bureau international

(43) Date de la publication internationale
22 septembre 2005 (22.09.2005)

PCT

(10) Numéro de publication internationale
WO 2005/088902 A2

(51) Classification internationale des brevets[7] : H04L 9/34

(21) Numéro de la demande internationale :
PCT/FR2005/000553

(22) Date de dépôt international : 8 mars 2005 (08.03.2005)

(25) Langue de dépôt : français

(26) Langue de publication : français

(30) Données relatives à la priorité :
0450463 8 mars 2004 (08.03.2004) FR

(71) Déposant *(pour tous les États désignés sauf US)* : **MEDIALIVE** [FR/FR]; 111, avenue Victor Hugo, F-75116 Paris (FR).

(72) Inventeurs; et

(75) Inventeurs/Déposants *(pour US seulement)* : **LECOMTE, Daniel** [FR/FR]; 157, rue de La Pompe, F-75116 Paris (FR). **CAPOROSSI, Jérôme** [FR/FR]; 7, rue du 8 mai 1945, Bât. F, F-92340 Bourg-La-Reine (FR). **PARAYRE-MITZOVA, Daniela** [FR/FR]; 88, rue Philippe de Girard, Bât. B, Appt 132, F-75018 Paris (FR).

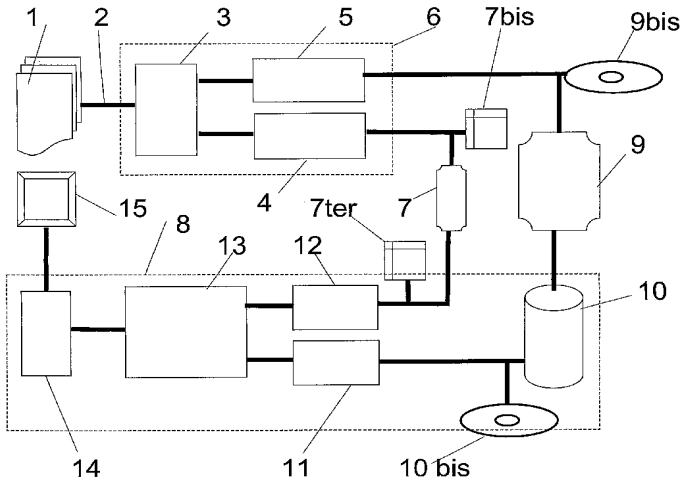
(74) Mandataire : **SAYETTA, Julien**; Breesé Derambure Majerowicz, 38, avenue de l'Opéra, F-75002 Paris (FR).

(81) États désignés *(sauf indication contraire, pour tout titre de protection nationale disponible)* : AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG,

[Suite sur la page suivante]

(54) Title: METHOD AND SYSTEM FOR THE SECURE DISTRIBUTION OF COMPRESSED DIGITAL TEXTS

(54) Titre : PROCEDE ET SYSTEME DE DISTRIBUTION SECURISEE DE TEXTES NUMERIQUES COMPRESSES



(57) Abstract: The invention relates to a method for the secure distribution of compressed digital texts comprising blocks of binary data and resulting from transformations applied to an original text. The inventive method comprises two steps, namely: a preparatory step consisting in modifying at least one binary datum in one of the aforementioned blocks using at least one substitution operation involving the extraction of said datum from a block and the replacement thereof with a decoy; and a step consisting in transmitting (i) a modified compressed digital text (5) that conforms to the format of the original text, comprising blocks that were modified during the preparatory step, and, over a separate channel from the modified compressed text (5), (ii) a piece of complementary digital information (4) which can be used to restore the original compressed digital text (1) on the destination equipment from the modified compressed digital text (5) and said complementary information (4). The invention also relates to a system which is used to implement said method.

(57) Abrégé : La présente invention se rapporte à un procédé pour la distribution sécurisée de textes numériques compressés formés de blocs de données binaires, issus de transformations appliquées à un texte original, et comporte deux étapes : une étape préparatoire consistant à modifier au moins une donnée binaire dans un desdits blocs selon au moins une opération

[Suite sur la page suivante]

WO 2005/088902 A2

MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) États désignés *(sauf indication contraire, pour tout titre de protection régionale disponible)* **:** ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), eurasien (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), européen (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Publiée :

— *sans rapport de recherche internationale, sera republiée dès réception de ce rapport*

En ce qui concerne les codes à deux lettres et autres abréviations, se référer aux "Notes explicatives relatives aux codes et abréviations" figurant au début de chaque numéro ordinaire de la Gazette du PCT.

de substitution consistant en l'extraction au sein d'un bloc de cette donnée binaire et son remplacement par un leurre, et une étape de transmission d'un texte numérique compressé modifié (5) conforme au format du texte original, constitué par des blocs modifiés au cours de l'étape préparatoire et par une voie séparée dudit texte numérique compressé modifié (5), d'une information complémentaire (4) numérique permettant de reconstituer le texte numérique compressé original (1), sur l'équipement destinataire, à partir dudit texte numérique compressé modifié (5) et de ladite information complémentaire (4). La présente invention concerne également un système pour la mise en oeuvre dudit procédé.

PROCÉDÉ ET SYSTÈME DE DISTRIBUTION SÉCURISÉE DE TEXTES NUMÉRIQUES COMPRESSÉS

La présente invention se rapporte au domaine des données binaires issues de transformations appliquées sur des textes numériques.

On se propose dans la présente invention de fournir un système permettant de protéger et de distribuer de manière sécurisée des textes numériques compressés et de restituer le texte numérique original, tout en prévenant une utilisation ou un accès non autorisé aux dits textes numériques compressés.

Dans la suite, on définit par le terme « texte » une succession de caractères issus d'un alphabet de lettres ou de chiffres, et de signes de ponctuation.

Dans la suite, on définit par le terme « texte numérique » la succession d'octets représentant des caractères issus d'un alphabet et/ou des signes de ponctuations et/ou des données de mise en forme et d'affichage d'un texte sur un écran de visualisation.

Dans la suite, on définit par le terme « texte numérique compressé » le flux de données binaires issu d'un algorithme de compression statistique appliqué au texte numérique.

Dans la suite, on définit l'action d'afficher un texte numérique compressé comme la série d'opérations consistant à lire et à décoder la succession de données binaires qui constituent le texte numérique compressé pour restituer le texte sur un écran de visualisation afin qu'il puisse être lu et compris d'un point de vue sémantique par un être humain.

La présente invention se rapporte plus particulièrement à un dispositif capable de transmettre de façon sécurisée un ensemble de textes numériques compressés vers un écran de visualisation et/ou pour être enregistrés

sur le disque dur d'un ordinateur ou sur le support d'enregistrement d'un boîtier reliant le réseau de télétransmission à l'écran de visualisation tel qu'un écran de télévision ou un moniteur d'ordinateur personnel, mais en évitant toute utilisation frauduleuse comme la possibilité de faire des copies illicites de contenus textuels ou de textes numériques compressés. L'invention concerne également un système client-serveur, entre le serveur qui fournit les textes numériques compressés sécurisés, et le client qui affiche, lit, enregistre ou imprime les textes numériques compressés.

Avec les solutions actuelles, il est possible de transmettre des textes et documents volumineux sous forme numérique via des réseaux de télécommunication type câble, DSL (Digital Subscriber Line) ou BLR (Boucle Locale Radio). Par ailleurs, pour éviter le piratage des œuvres et documents confidentiels ainsi diffusés, ces derniers sont souvent cryptés ou brouillés par divers moyens bien connus de l'homme de l'art.

Concernant la distribution sécurisée de textes et de données binaires, l'art antérieur connaît le document WO9805142 « Multi matrix encryption for private transmission of data », présentant un procédé et un équipement pour la protection de données et leur transmission sécurisée à travers un réseau électronique. Le document concerne le cryptage de données textuelles à l'aide de matrices de caractères ASCII générées par des clés. Les trois éléments clés en entrée sont un code PIN (Personal Identication Number), le numéro de compte bancaire de l'utilisateur et un mot de passe. Ces trois clés déclenchent la génération d'une matrice A et d'une matrice B. Les matrices A et B sont générées de manière pseudo aléatoire à l'aide d'une fonction analytique, comme une fonction logarithmique, une fonction trigonométrique, une fonction racine carrée ou autre. La distribution des caractères au sein des matrices A et B est

irrégulière et chaque caractère est unique. Avantageusement, trois valeurs de contrôle d'intégrité sont calculées et incorporées dans le flux protégé, dont une qui représente la somme des données textuelles en entrée, les deux autres étant relatives aux trois éléments clés en entrée. Les données en entrée sont transformées en valeur décimale de quatre chiffres, par des opérations de permutation, addition, soustraction, multiplication, division, et elles sont ensuite divisées en deux valeurs de deux chiffres. Ces deux valeurs sont indexées par rapport aux éléments des matrices A et B pour former le flux de données protégé. Toutefois, à cause de la division en deux parties pour l'indexation, la taille du flux protégé augmente considérablement par rapport à la taille des données initiales. De plus toutes les données protégées, ainsi que les trois valeurs de contrôle générées sont présentes à l'intérieur du flux protégé. Cet art antérieur ne correspond donc pas aux critères de haute sécurité, objectif de la présente invention.

La protection des textes numériques compressés, réalisée de façon conforme à la présente invention, est basée sur le principe de suppression et de remplacement de certaines informations codant les textes numériques compressés originaux par une méthode quelconque, soit : substitution, modification, permutation ou déplacement de l'information. Cette protection est également basée sur la connaissance de la structure des données binaires à la sortie de l'encodeur produisant les textes numériques compressés.

La présente invention concerne le principe général d'un procédé de sécurisation de textes numériques compressés. La solution consiste à extraire et à conserver en permanence et hors d'accès pour l'utilisateur, en fait dans le réseau de distribution, une partie du texte numérique compressé enregistré chez le client ou envoyé en

ligne, cette partie étant primordiale pour exploiter ledit texte numérique compressé sur un écran d'affichage, mais étant d'un volume très faible par rapport au volume total du texte numérique compressé enregistré chez l'usager ou reçu en ligne. La partie manquante sera transmise via le réseau de distribution au moment de l'exploitation dudit texte numérique compressé.

Le texte numérique compressé étant séparé en deux parties non égales, la plus grande partie du texte numérique compressé original séparé est appelée «texte numérique compressé modifié» et sera donc transmise via un réseau de diffusion classique large bande ou bande étroite, alors que la partie manquante appelée «information complémentaire» sera envoyée à la demande via un réseau de télécommunication bande étroite comme les réseaux téléphoniques classiques ou les réseaux cellulaires de type GSM, GPRS ou UMTS ou en utilisant une petite partie d'un réseau de type DSL ou BLR, ou en utilisant un sous-ensemble de la bande passante partagée sur un réseau câblé, ou encore via un support physique comme une carte à mémoire ou tout autre support. Avantageusement, les deux réseaux peuvent être confondus, tout en gardant les deux voies de transmission séparées. Le texte numérique compressé original est reconstitué sur l'équipement destinataire par un module de synthèse à partir du texte numérique compressé modifié et de l'information complémentaire.

Pour mette en œuvre le procédé, l'invention réalise un système de protection, comprenant un module d'analyse - protection et un module de recomposition basés sur un format numérique issu de l'encodage d'un texte numérique utilisant des algorithmes de compressions statistiques. Le module d'analyse et protection proposé par l'invention repose sur la substitution par des « leurres » ou sur la modification d'une partie des données binaires composant le texte numérique compressé original. Le fait d'avoir enlevé et

substitué une partie des données originales du texte numérique compressé original lors de la génération du texte numérique compressé modifié ne permet pas la recomposition dudit texte numérique compressé original à partir des seules données dudit texte numérique compressé modifié.

Se basant sur les caractéristiques des textes numériques compressés, plusieurs variantes du procédé de protection sont mises en œuvre et sont illustrées avec des exemples de réalisation.

La présente invention se rapporte plus particulièrement à un dispositif capable de transmettre de manière sécurisée un texte numérique vers un afficheur et/ou pour être enregistré dans la mémoire du dispositif de sauvegarde d'un boîtier reliant le réseau de télétransmission à l'afficheur, tout en préservant le contenu sémantique du texte, mais en évitant la possibilité que le texte numérique puisse être lu et copié de manière illicite.

Un texte numérique compressé généré par un algorithme de compression statistique à partir d'un texte numérique est constitué d'une succession de données binaires représentants des codes et/ou des entrées dans des tables de codage et/ou des pointeurs vers des positions au sein du texte numérique.

La présente invention consiste, après analyse du texte numérique compressé, à extraire au sein du texte numérique compressé au moins une donnée binaire originale représentant un code ou une entrée dans une table de codage ou un pointeur, cette donnée étant choisie de manière aléatoire, et à la remplacer par une donnée binaire appelée leurre, de même taille et de même nature mais de valeur aléatoire afin de générer un texte numérique compressé modifié conforme au format du texte numérique compressé original. L'affichage du texte numérique compressé modifié restitue alors un texte

illisible et/ou incompréhensible d'un point de vue sémantique pour un être humain.

Selon une variante de l'invention, la donnée binaire originale à extraire est choisie de manière déterministe.

Selon une autre variante de l'invention, la valeur de la donnée binaire leurre est calculée de manière déterministe.

Selon une autre variante de l'invention, la donnée binaire leurre est de taille différente de la taille de la donnée binaire originale.

L'invention selon son acception la plus générale concerne un procédé pour la distribution sécurisée de textes numériques compressés formés de blocs de données binaires, issus de transformations numériques appliquées à un texte original, caractérisé en ce qu'il comporte :

- une étape préparatoire consistant à modifier au moins une donnée binaire dans un desdits blocs selon au moins une opération de substitution consistant en l'extraction au sein d'un bloc de cette donnée binaire et son remplacement par un leurre,
- une étape de transmission :
  i. d'un texte numérique compressé modifié conforme au format du texte numérique compressé original, constitué par des blocs modifiés au cours de l'étape préparatoire et
  ii. par une voie séparée dudit texte numérique compressé modifié, d'une information complémentaire numérique permettant de reconstituer le texte numérique compressé original, à partir du calcul sur l'équipement destinataire, en fonction dudit texte numérique compressé modifié et de ladite information complémentaire.

Dans un mode de mise en œuvre, ladite donnée binaire représente une entrée dans une table de codage et le leurre représente une entrée différente dans ladite table de codage.

Dans un autre mode de mise en œuvre, la table de codage est construite de manière dynamique lors du décodage.

Avantageusement, la table de codage est prédéfinie par un standard donné ou une norme donnée.

Avantageusement, ladite donnée binaire représente une position antérieure au sein du texte numérique généré au cours du décodage et le leurre représente une position antérieure différente au sein dudit texte numérique généré au cours du décodage.

Dans un mode de mise en œuvre, le texte numérique compressé modifié (5) est conforme au standard du texte numérique compressé original (1).

Dans un autre mode de mise en œuvre, le texte numérique compressé modifié (5) est conforme au format du texte numérique compressé original (1).

Dans un mode de réalisation, ladite donnée binaire et ledit leurre sont de même taille.

Dans un autre mode de réalisation, ladite donnée binaire et ledit leurre sont de tailles différentes.

De préférence, la série de données binaires est codée de manière différentielle.

Dans une variante, le texte numérique compressé modifié est de même taille que le texte numérique compressé original.

Dans une autre variante, le texte numérique compressé modifié est de taille différente par rapport à celle du texte numérique compressé original.

De préférence, le texte numérique compressé reconstitué à partir du texte numérique compressé modifié est strictement identique au texte numérique compressé original.

Avantageusement, le procédé est appliqué à des textes numériques compressés issus du format de compression LZW.

Avantageusement, le procédé est appliqué à des textes numériques compressés issus du format de compression ZLIB/DEFLATE.

Avantageusement, le procédé est appliqué à des textes numériques compressés issus du format Adobe PDF.

Avantageusement, le procédé est appliqué à des images numériques compressés issus du format TIFF.

Avantageusement, le procédé est appliqué à des images numériques compressés issus du format GIF.

La présente invention concerne également un système pour la mise en œuvre du procédé, comportant au moins un serveur contenant des textes numériques compressés originaux et comportant un dispositif d'analyse du texte numérique compressé, un dispositif de séparation du texte numérique compressé original en un texte numérique compressé modifié et en une information complémentaire en fonction de ladite analyse, au moins un réseau de télécommunication pour la transmission et au moins un dispositif sur l'équipement destinataire pour la recomposition du texte numérique compressé original en fonction dudit texte numérique compressé modifié et de ladite information complémentaire.

La présente invention sera mieux comprise à l'aide des exemples de réalisation qui suivent et qui concernent des algorithmes de compression statistiques de textes numériques bien connus par l'homme de l'art.

L'algorithme de compression LZW (Lempel-Ziv-Welch) est un algorithme de compression statistique adaptatif à longueur variable, qui a été adopté notamment comme standard de compression dans les standards TIFF (Tagged Image File Format), GIF (Graphics Interchange Format) ou Adobe PDF (Portable Document Format). L'algorithme LZW compresse aussi bien des données binaires (flux d'octets) que visuelles (flux de pixels), ou encore les données d'un texte numérique.

Les données issues de l'algorithme de compression LZW consistent en une séquence de codes qui ont une longueur comprise entre 9 et 12 bits. Chaque code représente soit un simple caractère (c'est-à-dire un octet compris entre 0 et 255), un marqueur de réinitialisation de table (valeur 256), un marqueur « fin des données » (valeur 257) ou encore une entrée dans une table (valeur > 258), cette entrée étant associée à une séquence d'octets rencontrée précédemment dans le texte numérique à compresser. Initialement, et à l'encodage comme au décodage, les codes ont une longueur de 9 bits (valeur comprise entre 0 et 257) et la table est initialisée avec les 258 premières entrées (les 256 valeurs d'un octet + le marqueur de réinitialisation 256 + le marqueur fin de données 257). Au fur et à mesure que le processus d'encodage (ou de décodage) progresse, de nouveaux codes sont ajoutés à la table, associés chacun à des séquences d'octets de longueurs variables qui seront susceptibles d'apparaître de manière récurrente dans le texte numérique à compresser (ou à décompresser). À chaque fois qu'une séquence d'octets qui est déjà apparue réapparaît dans le texte numérique, le code correspondant à l'entrée de la table stockant cette même séquence est émis dans le texte numérique compressé. De même, à la décompression, les codes sont systématiquement remplacés par la séquence d'octets lue à l'entrée correspondante de la table et une nouvelle entrée est ajoutée à la table afin de

stocker la séquence formée à partir de la séquence décodée précédemment. Ainsi, la table est construite dynamiquement de la même manière à l'encodage comme au décodage. Lorsque la longueur binaire des codes n'est plus suffisante pour représenter une entrée dans la table, elle est augmentée de 1 : ainsi, dès que le nombre d'entrées de la table atteint 510, les codes sont codés sur 10 bits (et de la même manière lorsque le nombre d'entrée atteint 1022 (11 bits) et 2046 (12 bits)). Cependant les codes ne dépassent jamais une longueur de 12 bits (4095 entrées au maximum). Au sein d'un texte numérique compressé, le code 256 peut apparaître plusieurs fois : la table est alors réinitialisée et la longueur binaire des codes réinitialisée à 9.

Lors de l'opération de protection d'un texte numérique compressé issu de l'algorithme LZW, un algorithme lit le flux d'octets et construit dynamiquement la table de la même manière qu'un algorithme de décompression LZW.

L'opération de protection d'un texte numérique compressé LZW consiste à extraire, de manière aléatoire et/ou déterministe, dans la séquence un ou plusieurs (ce nombre étant déterminé de manière aléatoire ou calculé) codes originaux et à les remplacer par un ou plusieurs codes « leurres » valides, ces codes « leurres » valides pointant sur des entrées dans la table. Un code « leurre » est dit valide lorsque la nouvelle entrée pointée de la table existe et que cette entrée correspond à une séquence d'octets de longueur identique à celle pointée par le code original.

Le texte numérique décompressé à partir du texte numérique compressé modifié est de même taille que le texte numérique original. Le texte affiché à partir du texte numérique compressé modifié consiste en une succession aléatoire de caractères alphabétiques et de signes de ponctuations qui n'est pas intelligible pour un être humain.

Le format Adobe PDF (Portable Document Format) utilise l'algorithme de compression statistique LZW pour compresser

des objets de type texte numérique au sein d'un document encodé au format PDF. Un objet de type texte numérique représente un paragraphe, une ou plusieurs pages de texte, la légende d'une figure. Chaque objet de type texte numérique est codé de manière indépendante. Ainsi, la présente invention permet de protéger certains textes numériques afin de rendre le texte affiché illisible et/ou incompréhensible tout en laissant d'autres objets textes au sein du même document PDF lisibles et compréhensibles.

Avantageusement, la présente invention permet de protéger des objets de type figure et de type image numérique incorporés au sein d'un document texte et issus d'un algorithme de compression statistique en les rendant incohérents du point de vue de la perception visuelle humaine, tout en laissant des objets textes au sein du même document PDF lisibles et compréhensibles.

Avantageusement, la présente invention permet de protéger des images numériques au format TIFF et GIF en les rendant incohérentes de point de vue perception visuelle humaine.

L'algorithme de compression zlib/deflate est une combinaison de deux algorithmes de compression statistique : Huffman et LZ77 (Lempel-Ziv 77). Il est notamment utilisé pour compresser des objets de type texte numérique et/ou image numérique et/ou figure dans le format Adobe PDF.

L'algorithme d'Huffman consiste à remplacer une succession de symboles dans un flux original et issus d'un certain alphabet par une suite de codes de longueurs variables, chaque code se substituant à un symbole dans le flux compressé. L'algorithme commence par analyser le nombre et la fréquence des symboles apparaissant dans le flux original afin de construire un arbre de codage à partir duquel il associe à chaque symbole rencontré un code de longueur inversement proportionnelle à la fréquence d'apparition du symbole dans le flux original. La

compression consiste alors à remplacer chaque symbole par son code associé. L'algorithme de décompression a cependant besoin de l'arbre de codage pour décompresser le flux compressé. Pour l'algorithme zlib/deflate, une version modifiée d'Huffman est cependant utilisée : l'arbre de codage est construit en respectant des règles supplémentaires qui lui confèrent une propriété d'unicité et l'algorithme de décompression n'a plus besoin de l'arbre de codage mais seulement des longueurs des codes utilisés afin de reconstruire ce dernier.

L'algorithme LZ77 identifie les séquences de données récurrentes dans un flux au sein d'une fenêtre glissante de taille fixée. Lorsqu'une séquence étant déjà apparue est de nouveau détectée, elle est remplacée dans le flux compressé par deux nombres : une distance d et une longueur l. La distance indique à quel endroit dans la fenêtre cette même séquence débute et la longueur indique combien de données comporte la séquence identifiée. À la décompression, à chaque fois que l'algorithme rencontre un couple (d,l), il recopie dans le flux sortant la séquence de données de longueur l lue à partir de la position courante moins d.

L'algorithme de compression zlib/deflate utilise trois modes de compression : un mode « pas de compression » pour les données ayant déjà été compressées, un mode LZ77 + Huffman classique, les arbres de codage étant définis dans les spécifications de l'algorithme et un mode LZ77 + Huffman modifié. Les données sont découpées en blocs, chaque bloc étant codé de manière indépendante selon l'un des trois modes précédemment cités.

Dans les modes 2 et 3, les données sont codées tout d'abord selon LZ77 et une séquence de symboles est ainsi générée, ces symboles étant de type « caractère » (i.e. un octet dont la valeur est comprise entre 0 et 255) ou couple distance-longueur (d,l). Cette séquence de symboles est

alors compressée avec un algorithme d'Huffman classique (mode 2) ou un Huffman modifié (mode 3).

Conformément à l'invention, l'opération de protection d'un texte numérique compressé selon l'algorithme zlib/deflate consiste à modifier un ou plusieurs blocs codés selon les modes 2 ou 3. Les modifications consistent à extraire du texte numérique compressé un code d'Huffman codant un symbole de type « caractère » ou distance d et à le remplacer par un code d'Huffman valide. Un code d'Huffman est dit valide s'il est de même longueur que le code qu'il remplace et s'il correspond effectivement à un symbole codé de même type, c'est-à-dire caractère ou distance.

Un texte numérique compressé zlib/deflate modifié est de même taille que le texte numérique compressé zlib/deflate original. De même, le texte numérique décompressé à partir du texte numérique compressé modifié est de même taille que le texte numérique original.

L'affichage du texte numérique modifié produira un texte illisible et/ou incompréhensible pour tout être humain car affichant une succession de caractères et de signes de ponctuation sans aucune logique.

Les spécifications du format zlib définissent un champ de 4 octets ADLER32 situé à la fin des textes numériques compressés : ce champ stocke un identifiant unique du texte numérique original et il est utilisé lors de la décompression afin de vérifier l'intégrité du texte numérique. Dans le cas d'un texte numérique compressé zlib/deflate et modifié selon notre invention, la signature du texte numérique décompressé ne sera pas identique à celle du texte numérique original.

Avantageusement, la signature d'origine sera mise à jour lors de l'application de la protection.

On comprendra mieux l'invention à l'aide de la description, faite ci-après à titre purement explicatif, d'un mode de réalisation de l'invention, en référence à la figure 1 annexée qui illustre un mode de réalisation particulier du système permettant de protéger et de distribuer de manière sécurisée des textes numériques compressés selon l'invention.

Le texte numérique compressé (1) que l'on souhaite sécuriser est passé par le lien (2) à un module d'analyse - protection (3) qui va générer un texte numérique compressé modifié (5) au format identique au texte numérique compressé original (1) en dehors de ce que certaines des données binaires ont été remplacées par des valeurs différentes de celles d'origine, et est stocké dans le serveur (6). L'information complémentaire (4), de format quelconque, est également placée dans le serveur (6) et contient des informations relatives aux données du texte numérique compressé qui ont été modifiées, remplacées, substituées ou déplacées, et à leurs valeurs ou emplacements dans le texte numérique compressé original.

Avantageusement, le texte numérique compressé protégé (5) au format identique au texte numérique compressé original est ensuite transmis, via un réseau haut débit (9) de type hertzien, câble, satellite, ou autre réseau, au terminal de l'utilisateur (8), et plus précisément dans une mémoire (10).

Lorsque l'utilisateur (8) fait la demande d'affichage du texte présent dans la mémoire (10), deux éventualités sont possibles : soit l'utilisateur (8) ne possède pas tous les droits nécessaires pour exploiter le texte numérique compressé, dans ce cas, le texte numérique compressé modifié (5) généré par le module de protection (3) et présent sur la mémoire (10) est passé au système de synthèse (13), via une mémoire tampon de lecture (11), qui ne le modifie pas et le

transmet à l'identique à un afficheur capable de le décoder (14) et son contenu, détérioré par le module de protection (3) et incompréhensible de point de vue sémantique, est affiché sur l'écran de visualisation (15). Avantageusement, le texte numérique compressé modifié (5) généré par le module de protection (3) est passé directement via un réseau (9) à la mémoire tampon de lecture (11) puis au module de synthèse (13).

Soit le serveur (6) décide que l'utilisateur (8) possède les droits pour afficher correctement le texte numérique compressé. Dans ce cas, le module de synthèse (13) fait une demande d'affichage au serveur (6) contenant l'information complémentaire (4) nécessaire à la recomposition du texte numérique compressé original (1). Le serveur (6) envoie alors via le réseau de télécommunication (7) de type ligne téléphonique analogique ou numérique, DSL (Digital Subscriber Line) ou BLR (Boucle Locale Radio), via des réseaux DAB (Digital Audio Broadcasting) ou via des réseaux de télécommunications mobiles numériques (GSM, GPRS, UMTS), l'information complémentaire (4), permettant la reconstitution du texte numérique compressé, de façon à ce que l'utilisateur (8) puisse la stocker dans une mémoire tampon (12). Le module de synthèse (13) procède alors à la reconstitution du texte numérique compressé original à partir du texte numérique compressé modifié qu'il lit dans sa mémoire tampon de lecture (11), des champs modifiés dont il connaît les positions ainsi que les valeurs d'origine sont restituées grâce au contenu de l'information complémentaire lue dans la mémoire tampon (12) de recomposition. L'information complémentaire (4) qui est envoyée au module de recomposition est spécifique pour chaque utilisateur et dépend de ses droits, par exemple l'utilisation unique ou multiple, droit de faire une ou plusieurs copies privées, retard ou anticipation de paiement.

Avantageusement, le texte numérique compressé modifié (5) est passé directement via un réseau (9) à la mémoire tampon de lecture (11) puis au module de synthèse (13).

Avantageusement, le texte numérique compressé modifié (5) est enregistré sur un support physique comme un disque de type CD-ROM, DVD, disque dur, carte à mémoire flash (9bis). Le texte numérique compressé modifié (5) sera ensuite lu depuis le support physique (9bis) par le lecteur de disque (10bis) du boîtier (8) pour être transmis à la mémoire tampon de lecture (11) puis au module de recomposition (13).

Avantageusement, l'information complémentaire (4) est enregistrée sur un support physique (7bis) de format carte de crédit, constitué par une carte à puce ou une carte à mémoire flash. Cette carte (7bis) sera lue par le module (12) du dispositif (8) qui comprend un lecteur de carte (7ter).

Avantageusement, la carte (7bis) contient les applications et les algorithmes qui seront exécutés par le module de recomposition (13).

Avantageusement, le dispositif (8) est un dispositif autonome, portable et mobile.

REVENDICATIONS

1. Procédé pour la distribution sécurisée de textes numériques compressés formés de blocs de données binaires, issus de transformations numériques appliquées à un texte original, caractérisé en ce qu'il comporte :

- une étape préparatoire consistant à modifier au moins une donnée binaire dans un desdits blocs selon au moins une opération de substitution consistant en l'extraction au sein d'un bloc de cette donnée binaire et son remplacement par un leurre,

- une étape de transmission :

i. d'un texte numérique compressé modifié conforme au format du texte numérique compressé original (1), constitué par des blocs modifiés au cours de l'étape préparatoire et

ii. par une voie séparée dudit texte numérique compressé modifié (5), d'une information complémentaire (4) numérique permettant de reconstituer le texte numérique compressé original (1), à partir du calcul sur l'équipement destinataire, en fonction dudit texte numérique compressé modifié (5) et de ladite information complémentaire (4).

2. Procédé pour la distribution sécurisée de textes numériques compressés formés de blocs de données binaires selon la revendication 1, caractérisé en ce que ladite donnée binaire représente une entrée dans une table de codage et que le leurre représente une entrée différente dans ladite table de codage.

3. Procédé pour la distribution sécurisée de textes numériques compressés formés de blocs de données binaires selon les revendications 1 et 2, caractérisé en ce que la

table de codage est construite de manière dynamique lors du décodage.

4. Procédé pour la distribution sécurisée de textes numériques compressés formés de blocs de données binaires selon les revendications 1 et 2, caractérisé en ce que la table de codage est prédéfinie par un standard donné ou une norme donnée.

5. Procédé pour la distribution sécurisée de textes numériques compressés formés de blocs de données binaires selon la revendication 1, caractérisé en ce que ladite donnée binaire représente une position antérieure au sein du texte numérique généré au cours du décodage et que le leurre représente une position antérieure différente au sein dudit texte numérique généré au cours du décodage.

6. Procédé pour la distribution sécurisée de textes numériques compressés formés de blocs de données binaires selon l'une des revendications précédentes, caractérisé en ce que ladite donnée binaire et ledit leurre sont de même taille.

7. Procédé pour la distribution sécurisée de textes numériques compressés formés de blocs de données binaires selon l'une des revendications précédentes, caractérisé en ce que ladite donnée binaire et ledit leurre sont de tailles différentes.

8. Procédé pour la distribution sécurisée de textes numériques compressés formés de blocs de données binaires selon l'une des revendications précédentes, caractérisé en ce que la série de données binaires est codée de manière différentielle.

9. Procédé pour la distribution sécurisée de textes numériques compressés formés de blocs de données binaires selon l'une des revendications précédentes, caractérisé en ce que le texte numérique compressé modifié (5) est conforme au standard du texte numérique compressé original (1).

10. Procédé pour la distribution sécurisée de textes numériques compressés formés de blocs de données binaires selon l'une des revendications précédentes, caractérisé en ce que le texte numérique compressé modifié (5) est conforme au format du texte numérique compressé original (1).

11. Procédé pour la distribution sécurisée de textes numériques compressés formés de blocs de données binaires selon l'une des revendications précédentes, caractérisé en ce que le texte numérique compressé modifié (5) est de même taille que le texte numérique compressé original (1).

12. Procédé pour la distribution sécurisée de textes numériques compressés formés de blocs de données binaires selon l'une des revendications précédentes, caractérisé en ce que le texte numérique compressé modifié (5) est de taille différente par rapport à celle du texte numérique compressé original (1).

13. Procédé pour la distribution sécurisée de textes numériques compressés formés de blocs de données binaires selon une des revendications précédentes, caractérisé en ce que le texte numérique compressé reconstitué à partir du texte numérique compressé modifié (5) est strictement identique au texte numérique compressé original (1).

14. Procédé pour la distribution sécurisée de textes numériques compressés formés de blocs de données binaires selon l'une des revendications précédentes, caractérisé en

9. Procédé pour la distribution sécurisée de textes numériques compressés formés de blocs de données binaires selon l'une des revendications précédentes, caractérisé en ce que le texte numérique compressé modifié (5) est conforme au standard du texte numérique compressé original (1).

10. Procédé pour la distribution sécurisée de textes numériques compressés formés de blocs de données binaires selon l'une des revendications précédentes, caractérisé en ce que le texte numérique compressé modifié (5) est conforme au format du texte numérique compressé original (1).

11. Procédé pour la distribution sécurisée de textes numériques compressés formés de blocs de données binaires selon l'une des revendications précédentes, caractérisé en ce que le texte numérique compressé modifié (5) est de même taille que le texte numérique compressé original (1).

12. Procédé pour la distribution sécurisée de textes numériques compressés formés de blocs de données binaires selon l'une des revendications précédentes, caractérisé en ce que le texte numérique compressé modifié (5) est de taille différente par rapport à celle du texte numérique compressé original (1).

13. Procédé pour la distribution sécurisée de textes numériques compressés formés de blocs de données binaires selon une des revendications précédentes, caractérisé en ce que le texte numérique compressé reconstitué à partir du texte numérique compressé modifié (5) est strictement identique au texte numérique compressé original (1).

14. Procédé pour la distribution sécurisée de textes numériques compressés formés de blocs de données binaires selon l'une des revendications précédentes, caractérisé en

ce qu'il est appliqué à des textes numériques compressés issus du format de compression LZW.

15. Procédé pour la distribution sécurisée de textes numériques compressés formés de blocs de données binaires selon l'une des revendications précédentes, caractérisé en ce qu'il est appliqué à des textes numériques compressés issus du format de compression ZLIB/DEFLATE.

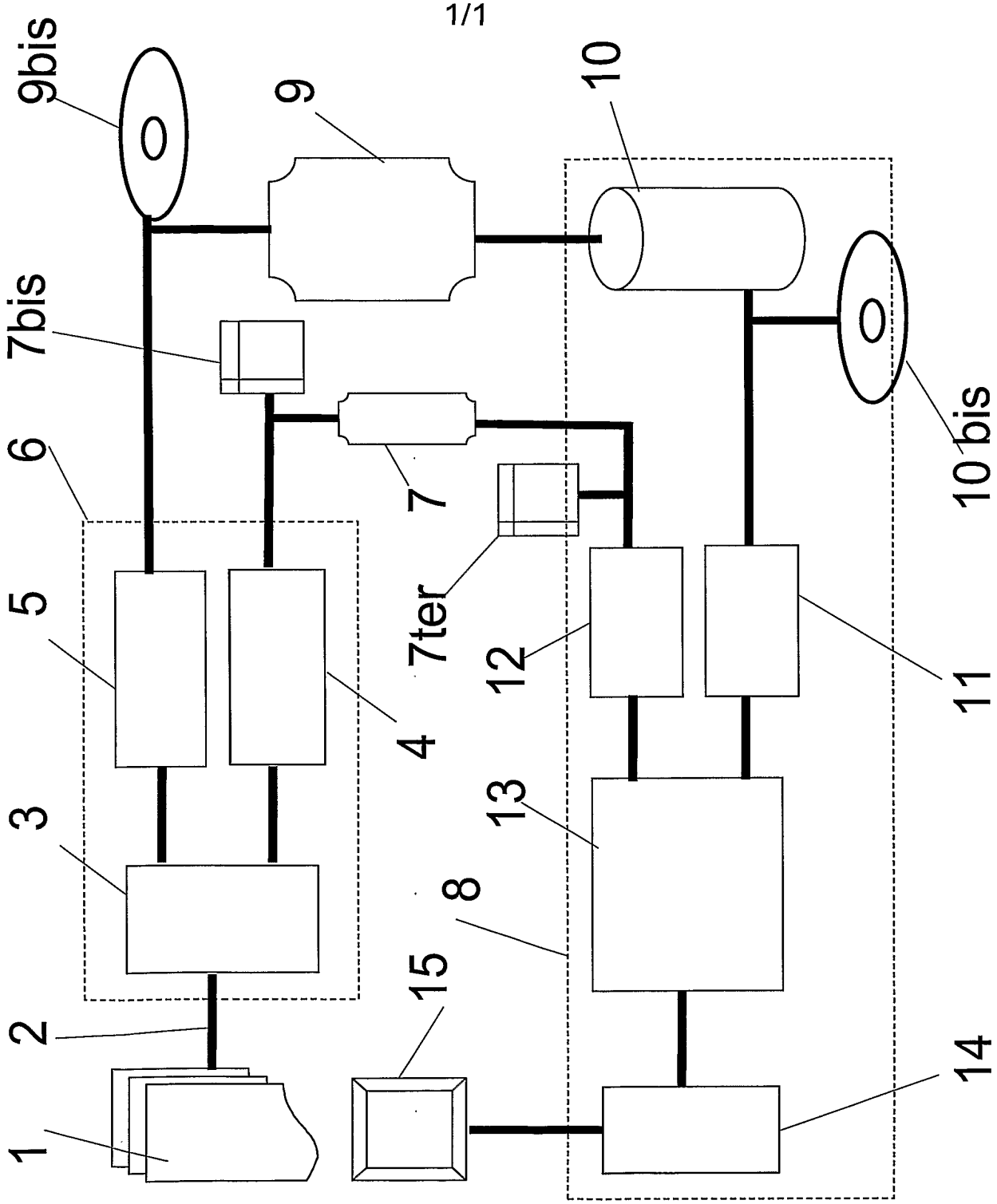
16. Procédé pour la distribution sécurisée de textes numériques compressés formés de blocs de données binaires selon l'une des revendications précédentes, caractérisé en ce qu'il est appliqué à des textes numériques compressés issus du format Adobe PDF.

17. Procédé pour la distribution sécurisée de textes numériques compressés formés de blocs de données binaires selon l'une des revendications précédentes, caractérisé en ce qu'il est appliqué à des images numériques compressées issues du format TIFF.

18. Procédé pour la distribution sécurisée de textes numériques compressés formés de blocs de données binaires selon l'une des revendications précédentes, caractérisé en ce qu'il est appliqué à des images numériques compressées issues du format GIF.

19. Système pour la mise en œuvre du procédé selon l'une des revendications précédentes, comportant au moins un serveur contenant des textes numériques compressés originaux et caractérisé en ce qu'il comporte un dispositif d'analyse du texte numérique compressé, un dispositif de séparation du texte numérique compressé original (1) en un texte numérique compressé modifié (5) et en une information complémentaire (4) en fonction de ladite analyse, au moins un réseau de

télécommunication pour la transmission et au moins un dispositif sur l'équipement destinataire pour la recomposition du texte numérique compressé original (1) en fonction dudit texte numérique compressé modifié (5) et de ladite information complémentaire (4).


9bis
9
10
7bis
10bis
6
7
7ter
5
12
11
4
3
13
8
15
2
14
1